SCIENCE CHINA
Life Sciences

• LETTER TO THE EDITOR •

July 2015 Vol.58 No.7: 724–726
doi: 10.1007/s11427-015-4888-2

Identification of a novel HIV-1 intra-circulating recombinant form 01_AE in China: a descendant of the previously identified CRF01_AE transmission clusters 1 and 6

NING ChuanYi[1,3†], LI XingGuang[2†], TANG WeiMing[1], ZHOU Bo[3], CAI WeiPing[1*] & TUCKER Joseph D.[1*]

[1]*Guangzhou Eighth People's Hospital & University of North Carolina and China Project, Guangzhou 510060, China;*
[2]*HIV Molecular Epidemiology and Virology Research Group, State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China;*
[3]*Research Center for Regenerative Medicine, Guangxi Medical University, Nanning 530021, China*

Received April 16, 2015; accepted May 20, 2015; published online June 19, 2015

Citation: Ning CY, Li XG, Tang WM, Zhou B, Cai WP, Tucker JD. Identification of a novel HIV-1 intra-circulating recombinant form 01_AE in China: a descendant of the previously identified CRF01_AE transmission clusters 1 and 6. Sci China Life Sci, 2015, 58: 724−726, doi: 10.1007/s11427-015-4888-2

Dear Editors,

Here, we report a novel HIV-1 intra-circulating recombinant form 01_AE (intra-CRF01_AE: CRF01-1AE/ CRF01-6AE) composed of CRF01_AE transmission clusters 1 and 6, which was identified in a heterosexual male from Fujian, with one breakpoint observed in the *vif* gene. The CRF01_AE region I (HXB2: 868–5184) of the recombinant clustered with CRF01_AE transmission cluster 1, which is mainly circulating among injection drug users (IDUs) and heterosexuals in southern and southwestern China, with the support of a 100% bootstrap value. The CRF01_AE region II (HXB2: 5185–9605) of the recombinant clustered with CRF01_AE transmission cluster 6, which mainly prevails among heterosexuals in southern China, with the support of a 100% bootstrap value. To our knowledge, this is the first detection of a novel HIV-1 intra-CRF01_AE recombinant form (CRF01-1_AE/CRF01-6_ AE) in Fujian, China, which indicates ongoing active HIV-1 transmission among heterosexuals in this region. This identification will be helpful to illustrate the genetic diversity of CRF01_AE and to better understand HIV-1 epidemiology, pathogenesis, and vaccine development.

The HIV-1 circulating recombinant form (CRF) 01_AE represents a putative recombinant form of HIV-1 subtypes A and E, although the subtype E lineage has not been found. CRF01_AE, which originated from Central Africa, was first identified among female sex workers in northern Thailand in 1989, and has now become the most prevalent clade in Thailand and neighboring countries in Southeast Asia [1,2]. In the early 1990s, CRF01_AE was first identified among the heterosexual population and IDUs in the Yunnan and Guangxi provinces of China, and then spread to the eastern coastal areas. Since then, CRF01_AE has been the most common HIV-1 strain identified among newly reported cases in China [3,4]. Overall, there are at least seven distinct transmission clusters of the CRF01_AE lineages, which are widely distributed among several risk populations in China [3]. This may increase the complexity of the HIV epidemic and could further complicate the design of a HIV vaccine. Our prior knowledge indicates that there may be several intra-CRF01_AE recombinant viruses between the distinct CRF01_AE transmission clusters, indicating the necessity to identify these clusters. Here we report a novel HIV-1 intra-CRF01_AE recombinant form (CRF01-1AE/ CRF01-6AE), which should help to better understand the diversity of the CRF01_AE genotypes.

†Contributed equally to this work
*Corresponding anthor (email: caiwp@163.net; jdtucker@med.unc.edu)

We extracted the complete genome or near full-length genome (NFLG) sequences of CRF01_AE strains published in the Los Alamos National Laboratory (LANL) HIV database (http://www.hiv.lanl.gov) by using the sequence search tool and focusing on the genomic region of the complete genome. As a result, we identified an intra-CRF01_AE recombinant strain (Fj056, GenBank accession number: EF036529) from a total of 346 complete whole genome or NFLG sequences of CRF01_AE strains. Fj056 was first detected from a 25-year-old heterosexual transmitted male from Fuzhou city in southeast China in 2005, where the most common transmission route was sexual transmission and the most common strains were CRF01_AE strains. The NFLG sequence of Fj056 was aligned against all known HIV-1 group M reference sequences representing subtypes or sub-subtypes (A1, A2, B, C, D, F1, F2, G, H, J and K), and aligned with all complete genome or NFLG sequences of CRF01_AE obtained from the LANL HIV sequence database. Alignment was conducted by using the online software Gene Cutter in LANL. We further selected three to four NFLG sequences from each well-supported distinct phylogenetic CRF01_AE transmission cluster identified in China. We also selected three known NFLG sequences from Central Africa. Overall, a total of 25 NFLG sequences of CRF01_AE lineages from three population groups (IDUs, heterosexuals and men who have sex with men [MSM]) in 10 provinces across China were collected during 2002–2010.

The identified novel recombinants had a different genomic structure from all recombinant forms reported previously. The results of phylogenetic analysis indicated that the genome sequence of Fj056 shared a more recent common ancestor with CRF01_AE transmission clusters 1 and 6, but was branched independently from these clusters. These results indicated that Fj056 might be a new recombinant arising from the CRF01_AE transmission clusters 1 and 6 (Figure S1). SimPlot analysis using the same reference sequence dataset (excluding the three CRF01_AE sequences from Central Africa) also showed that the genome sequence of Fj056 was composed of gene regions of CRF01_AE transmission clusters 1 and 6 (Figure 1a). BootScan analysis revealed that the breakpoint corresponded to HXB2 nucleotide position 5184 (Figure 1b). Similar results were obtained by using the RDP4 software package. Subregion tree analyses further identified the parental origin of each region in the recombinant genome: region I (HXB2: 868–5184)= CRF01_AE transmission cluster 1 and region II (HXB2: 5185–9605)=CRF01_AE transmission cluster 6 (Figure S2).

According to early reports, HIV-1 CRF01_AE was identified as the major subtype in China [4]. However, increasing numbers of full-length HIV-1 CRF01_AE genomic sequences have recently been reported in China. Fj056 represents a recent genetic characterization of HIV-1 CRF01_AE strains in Fujian province. Determination of the full-length genomic sequences of the intra-CRF01_AE recombinant form (CRF01-1AE/CRF01-6AE) may help to illustrate the genetic diversity of CRF01_AE and contribute to our understanding of HIV-1 epidemiology.

The official nomenclature of HIV-1 includes groups (currently M, N, O, and P) and subtypes within the HIV-1 M group. The so-called “circulating recombinant forms” include viruses such as CRF01_AE, CRF02_AG, and CRF04_cpx, which are apparently ancient recombinants with some parts of their genomes being more similar to one subtype, and other parts being more similar to other subtype(s). In addition, there are intra-subtype recombinants. The official nomenclature also includes some so-called “sub-subtypes” such as A1, A2, F1, and F2. There are also many unofficial designations of local strains and subclades within different subtypes, such as the “B-prime” and “Thai-

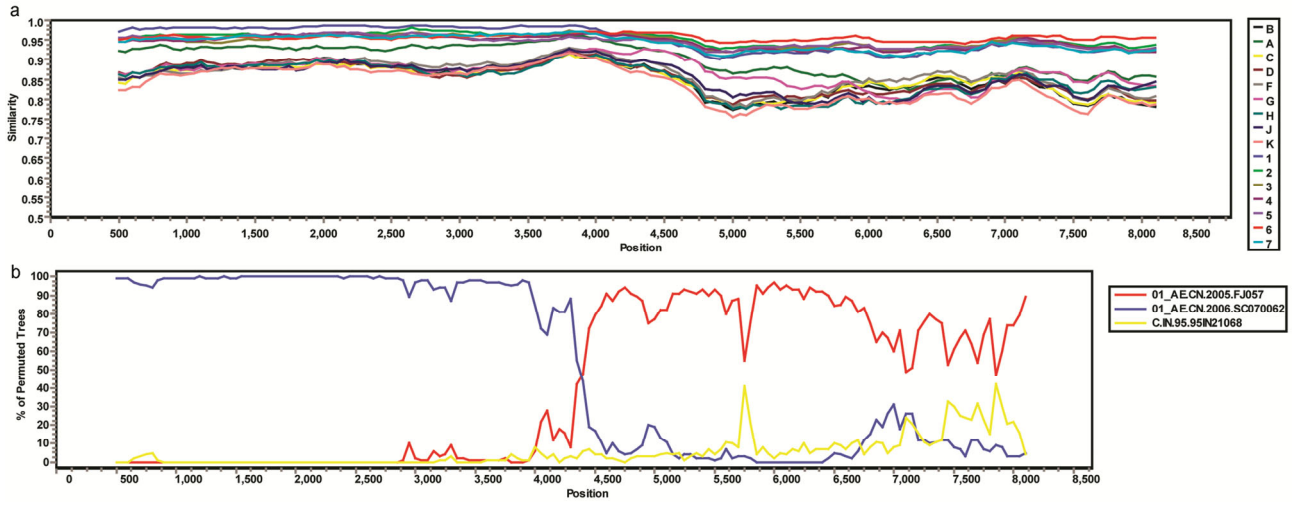

Figure 1 Recombinant analyses of the novel intra-CRF01_AE recombinant strain Fj056. a, SimPlot analysis of Fj056. The analysis was performed with a sliding window of 1,000 bp and a step size of 50 bp. b, BootScan analysis of Fj056. The reference strains were CRF01_AE transmission cluster 6, FJ057 (red), CRF01_AE transmission cluster 1, SC070062 (blue), and subtype C, 95IN21068 (yellow). The BootScan window was 1,000 bp with a step of 50 bp.

B" clades, and the "A3" and "A4" viruses. Recombination identification tools for HIV-1 M group viruses were specifically designed to detect inter-subtype recombinants, but most of these can be modified to detect intra-subtype recombination. The LANL HIV Databases Recombination Identification Program (RIP) tool (http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html) was designed to allow users to input reference sequences, including the seven clades of interest (01-1AE–01-7AE) of this study. However, after the subtype clades were identified, assigned them unofficial names is a slippery slope that could lead to misunderstanding of these subtype clades. Therefore, there is still no clear criterion for defining one sub-subtype worldwide, and we are not confident whether an intra-subtype referred to in one study in fact represent the same clade as reported in another study. For example, in 2013, Kondo et al. [5] reported at least two distinct clades of CRF01_AE among the Chinese MSM population, and it is unclear as to whether these are the same clades reported by Feng et al. [3]. Thus, it is critical to extend the official nomenclature to clearly characterize HIV sub-subtype.

In summary, we found a novel HIV-1 intra-CRF01_AE recombinant strain Fj056, which is composed of gene regions from the CRF01_AE transmission clusters 1 and 6 in Fujian province. There are currently at least seven distinct transmission clusters of CRF01_AE lineages circulating in China, and we expect that other novel HIV-1 intra-CRF01_AE recombinant strains could be detected with further screening.

The authors declare that they have no conflict of interest.

The authors thank the South China-UNC STD Research Training Center and Postdoctoral Science Foundation funded project of Guangxi Medical University. This work was supported by National Insitutes of Health Fogarty International Center grant (1D43TW009532-01), and Chinese National Key Technologies R&D Program for the 12th Five-year Plan (2012ZX10001003-003).

1 Gao F, Robertson DL, Morrison SG, Hui H, Craig S, Decker J, Fultz PN, Girard M, Shaw GM, Hahn BH, Sharp PM. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. J Virol, 1996, 70: 7013–7029

2 Nelson KE, Celentano DD, Suprasert S, Wright N, Eiumtrakul S, Tulvatana S, Matanasarawoot A, Akarasewi P, Kuntolbutra S, Romyen S, Sirisopana N, Theetranont C. Risk factors for HIV infection among young adult men in northern Thailand. JAMA, 1993, 270: 955–960

3 Feng Y, He X, Hsi JH, Li F, Li X, Wang Q, Ruan Y, Xing H, Lam TT, Pybus OG, Takebe Y, Shao Y. The rapidly expanding CRF01_AE epidemic in China is driven by multiple lineages of HIV-1 viruses introduced in the 1990s. AIDS, 2013, 27: 1793–1802

4 Huang HL, Zheng J, Yan PP, Cheng L, Lin X, Zhang ZS, Yan YS. Genetic characterization of CRF01_AE full-length human immunodeficiency virus type 1 sequences from Fujian, China. AIDS Res Hum Retroviruses, 2007, 23: 569–574

5 Kondo M, Lemey P, Sano T, Itoda I, Yoshimura Y, Sagara H, Tachikawa N, Yamanaka K, Iwamuro S, Matano T, Imai M, Kato S, Takebe Y. Emergence in Japan of an HIV-1 variant associated with transmission among men who have sex with men (MSM) in China: first indication of the International Dissemination of the Chinese MSM lineage. J Virol, 2013, 87: 5351–5361

Supporting Information

Figure S1 Phylogenetic analyses of the NFLG nucleotide sequence of the novel intra-CRF01_AE recombinant strain Fj056. The neighbor-joining tree represents the relationship of Fj056 with reference sequences. The solid circle (●) indicates Fj056. The seven unique CRF01_AE lineages detected in China are labeled 01-1AE through 01-7AE. Bootstrap analysis was performed with 1000 replicates and only bootstrap values ⩾70% are shown at the corresponding nodes. The scale bar represents a 2% genetic distance.

Figure S2 Phylogenetic trees for genome regions (I-II) of Fj056. Using the same methods as described in the legend to Figure S1, phylogenetic analysis was performed based on the recombinant breakpoint shown in Figure 1b.

The supporting information is available online at life.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.